

# Aggregating time preferences with decreasing impatience\*

Nina Anchugina<sup>1</sup>, Matthew Ryan<sup>2</sup>, and Arkadii Slinko<sup>1</sup>

<sup>1</sup>Department of Mathematics, University of Auckland

<sup>2</sup>School of Economics, Auckland University of Technology

n.anchugina@auckland.ac.nz, matthew.ryan@aut.ac.nz, a.slinko@auckland.ac.nz

April 2016

**Abstract.** It is well-known that for a group of time-consistent decision makers their collective time preferences may become time-inconsistent. Jackson and Yariv [11] demonstrated that the result of aggregation of exponential discount functions always exhibits present bias. We show that when preferences satisfy the axioms of Fishburn and Rubinstein [7], present bias is equivalent to decreasing impatience (DI). Applying the notion of comparative DI introduced by Prelec [19], we generalize the result of Jackson and Yariv [11]. We prove that the aggregation of distinct discount functions from comparable DI classes results in the collective discount function which is strictly more DI than the least DI of the functions being aggregated.

We also prove an analogue of Weitzman's [27] result, for hyperbolic rather than exponential discount functions. We show that if a decision maker is uncertain about her hyperbolic discount rate, then long-term costs and benefits will be discounted at a rate which is the probability-weighted harmonic mean of the possible hyperbolic discount rates.

**Keywords:** Discounting, hyperbolic discounting, decreasing impatience, aggregation.

**JEL Classification:** D71, D90.

\* We thank Matthew Jackson, Simon Grant and several seminar audiences for comments and suggestions. Arkadii Slinko was supported by the Marsden Fund grant UOA 1420, and Nina Anchugina gratefully acknowledges financial support from the University of Auckland.

# 1 Introduction

Sometimes decisions about timed outcomes have to be made by a group of individuals, such as boards, committees or households. It is natural to think that individuals may differ in the discounting procedure that they use. If the decision is to be made by a group of individuals it is desirable to have an aggregating procedure that suitably reflects the time preferences of all members. The natural option is to average the discount functions across individuals, which is equivalent to averaging the discounted utilities in the case when all agents have identical utility functions. This approach has been widely used in the existing literature on time preferences. It is known that such collective discount functions need not share properties that are common to the individual discount functions being aggregated. As Jackson and Yariv demonstrate [11], if individuals discount the future exponentially and there is a heterogeneity in discount factors, then their aggregate discount function exhibits present bias, which means that delaying two different dated-outcomes by the same amount of time can reverse the ranking of these outcomes. Moreover, when the number of individuals grows, in the limit the group discount function becomes hyperbolic [11].

Jackson and Yariv [11] give the following example of present-biased group preferences for a household with two time-consistent individuals, Constantine and Patience. Both have identical instantaneous utility functions, and discount the future exponentially, but Constantine has a discount factor of 0.5, whereas Patience has a discount factor of 0.8. Suppose that they need to choose between 10 utiles for each today or 15 utiles for each tomorrow. They calculate the aggregate discounted utility for each option:  $10 + 10 = 20$  and  $15(0.8 + 0.5) = 19.5$ . Therefore, 10 utiles today is chosen. Now suppose that they must choose between 10 utiles at time  $t \geq 1$  and 15 utiles at  $t + 1$ . The aggregate discounted utilities in this case are  $10(0.8^t + 0.5^t)$  and  $15(0.8^{t+1} + 0.5^{t+1})$ , respectively. For any  $t \geq 1$  the 15 utiles at  $t + 1$  is preferable to the 10 utiles at  $t$ , which reverses the initial preference for 10 utiles at  $t = 0$  over 15 utiles at  $t = 1$ . The behaviour of the household is present-biased.

Another scenario in which the aggregation of time preferences may be required is when a single decision maker is uncertain about the appropriate discount function to apply. For example, discounting may be affected by a survival function with a constant but uncertain hazard rate. Such scenarios are considered by Weitzman [27] and Sozou [24]. If the decision-maker maximizes expected discounted utility, then she maximizes discounted utility for a certainty equivalent discount function, calculated as the probability-weighted average of the different possible discount functions that may apply. Weitzman [27] shows that if each of the possible rates of time preference converges to some non-negative value (as time goes to infinity), then the certainty equivalent time preference function converges to the lowest of these limits. Similarly, Sozou [24] considers a decision maker whose discounting reflects a survival function with a constant, but uncertain, hazard rate. If this hazard rate is exponentially distributed, Sozou shows that the decision-maker's expected discount function is hyperbolic.

Of course, present bias is not limited to aggregate or expected discount functions. It is often observed in experiments that individual decision makers become decreasingly im-

patient (increasingly patient) as rewards are shifted further into the future. If a decision-maker is indifferent between an early outcome and a larger, later outcome, then delaying both outcomes by the same amount of time will often result in the larger, later outcome being preferred. Such subjects exhibit present bias, or strictly decreasing impatience (DI). Exponential discounting implies constant impatience, so it cannot explain *strictly* decreasing impatience, either globally or locally. The necessity of accommodating DI in individual time preference has made hyperbolic discounting a significant tool in behavioural economics. Several types of hyperbolic discount functions have been introduced, including quasi-hyperbolic discounting [17, 14], discounting for delay [1], proportional hyperbolic discounting [10, 16], and generalized hyperbolic discounting [15, 2]. Given the widespread use of hyperbolic discount functions to describe individual time preferences, it is important to understand the behaviour of aggregated, or averaged, hyperbolic functions.

The goal of this paper is twofold. Firstly, we seek to extend Jackson and Yariv’s result on the aggregation of exponential discount functions. Two individuals may differ in the rate at which their impatience decreases, but their respective levels of DI may be comparable – the preferences of the one may exhibit unambiguously more DI than the preferences of the other. As Prelec [19] proved, one individual exhibits more DI than another if the logarithm of the discount function of the former is more convex than that of the latter. Can we say anything about the level of DI of the weighted average of individual discount functions that can be (weakly) ordered by DI? Theorem 1 establishes that the weighted average always exhibits *strictly* more DI than the component with the *least* DI. This generalizes Jackson and Yariv’s result. Proposition 1 in [11] shows that the weighted average of exponential discount functions with different discount factors exhibits present bias. We show that when preferences satisfy the axioms of Fishburn and Rubinstein [7], Jackson and Yariv’s definition of present bias is equivalent to strictly decreasing impatience. Since all exponential discount functions exhibit constant impatience – they all exhibit the same degree of DI – Proposition 1 of Jackson and Yariv is a special case of our Theorem 1.

Our second goal is to prove an analogue of Weitzman’s [27] result: one in which discounting is hyperbolic but there is an uncertainty about the hyperbolic discount factor. The answer, given in Theorem 3, is very different to Weitzman’s answer for the case of exponential discounting. We show that the certainty equivalent hyperbolic discount factor converges, not to the lowest individual hyperbolic discount factor, but to the probability-weighted harmonic mean of the individual hyperbolic discount factors.

## 2 Preliminaries

In this section we introduce the framework for our investigation and define the two key concepts used in this paper: present bias and strictly decreasing impatience of preferences. We prove that these two concepts coincide when the Fishburn-Rubinstein axioms for a discounted utility representation are satisfied. Taking our lead from Pratt [18] and Arrow [3], these concepts are discussed in terms of log-convexity of discount functions, hence we introduce necessary results and definitions in this regard. Most results are known but

included to keep the paper self-contained.

## 2.1 Convexity and log-convexity

Convexity and log-convexity play an important role in the theory of discounting. Let  $I$  be an interval (finite or infinite) of real numbers. A function  $f: I \rightarrow \mathbb{R}$  is *convex* if for any two points  $x, y \in I$  and any  $\lambda \in [0, 1]$  it holds that:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

A function  $f$  is *strictly convex* if

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

for any  $x, y \in I$  such that  $x \neq y$  and any  $\lambda \in (0, 1)$ . If  $f$  is twice differentiable convexity is equivalent to  $f'' \geq 0$ , and strict convexity is equivalent to two conditions: the function  $f''$  is nonnegative on  $I$  and the set  $\{x \in I \mid f''(x) = 0\}$  contains no non-trivial interval [25].

The following equivalent definition of a (strictly) convex function is well known. A function  $f: I \rightarrow \mathbb{R}$  is (strictly) convex if for every  $x, y, v, z \in I$  such that  $x - y = v - z > 0$  and  $y > z$  we have

$$f(x) - f(y) \leq [ < ] f(v) - f(z).$$

Convexity is preserved under composition of functions, as shown in the following lemma, whose straightforward proof is omitted:

**Lemma 1.** *Let  $f_1: I \rightarrow \mathbb{R}$  be a non-decreasing and convex function and  $f_2: I \rightarrow \mathbb{R}$  be a convex function, such that the range of  $f_2$  is contained in the domain of  $f_1$ . Then the composition  $f = f_1 \circ f_2$  is a convex function. If, in addition,  $f_1$  is strictly increasing, and either  $f_1$  or  $f_2$  is strictly convex, then  $f$  is also strictly convex.*

A function  $f: I \rightarrow \mathbb{R}$  is called *log-convex* if  $f(x) > 0$  for all  $x \in I$  and  $\ln(f)$  is convex. It is called *strictly log-convex* if  $\ln(f)$  is strictly convex. It follows that if  $f$  is a (strictly positive) twice differentiable function, then log-convexity of  $f$  is equivalent to the condition  $f''f - (f')^2 \geq 0$ , while strict log-convexity of  $f$  requires, in addition, that the set

$$\{x \in I \mid f''(x)f(x) - [f'(x)]^2 = 0\}$$

contains no non-trivial interval. Log-convexity can also be expressed without using logarithms [5]. A function  $f: I \rightarrow \mathbb{R}$  is log-convex if and only if  $f(x) > 0$  for all  $x \in I$  and for all  $x, y \in I$  and  $\lambda \in [0, 1]$  we have:

$$f(\lambda x + (1 - \lambda)y) \leq f(x)^\lambda f(y)^{1-\lambda}. \quad (1)$$

The function  $f$  is strictly log-convex if inequality (1) is strict when  $x \neq y$  and  $\lambda \in (0, 1)$ .

The following result appears to be well known, but a formal reference is elusive so we have included a proof here for completeness.

**Lemma 2.** Let  $f, g: I \rightarrow \mathbb{R}$  be functions with  $f$  strictly log-convex and  $g$  log-convex. Then the sum  $f + g$  is strictly log-convex.

*Proof.* Since  $f(x) > 0$  and  $g(x) > 0$  for all  $x \in I$ , we have  $(f + g)(x) > 0$  for all  $x \in I$ . Let  $x, y \in I$  such that  $x \neq y$  and let  $\lambda \in (0, 1)$ . We must show that

$$f(\lambda x + (1 - \lambda)y) + g(\lambda x + (1 - \lambda)y) < (f(x) + g(x))^\lambda (f(y) + g(y))^{1-\lambda}.$$

Since  $f$  is strictly log-convex, we have

$$f(\lambda x + (1 - \lambda)y) < f(x)^\lambda f(y)^{1-\lambda}. \quad (2)$$

Analogously, since  $g(x)$  is log-convex:

$$g(\lambda x + (1 - \lambda)y) \leq g(x)^\lambda g(y)^{1-\lambda}. \quad (3)$$

Summing (2) and (3) we obtain:

$$f(\lambda x + (1 - \lambda)y) + g(\lambda x + (1 - \lambda)y) < f(x)^\lambda f(y)^{1-\lambda} + g(x)^\lambda g(y)^{1-\lambda}.$$

Denote  $a = f(x), b = f(y), c = g(x), d = g(y)$ . Note that  $a, b, c, d > 0$ . To prove the claim of the lemma, it is sufficient to show that:

$$a^\lambda b^{1-\lambda} + c^\lambda d^{1-\lambda} \leq (a + c)^\lambda (b + d)^{1-\lambda}. \quad (4)$$

Since  $(a + c)^\lambda (b + d)^{1-\lambda} > 0$  we can divide both parts of (4) by this expression to get

$$\left(\frac{a}{a + c}\right)^\lambda \left(\frac{b}{b + d}\right)^{1-\lambda} + \left(\frac{c}{a + c}\right)^\lambda \left(\frac{d}{b + d}\right)^{1-\lambda} \leq 1.$$

By the Weighted AM-GM inequality [6, Theorem 7.6, p. 74]:

$$\left(\frac{a}{a + c}\right)^\lambda \left(\frac{b}{b + d}\right)^{1-\lambda} \leq \lambda \frac{a}{a + c} + (1 - \lambda) \frac{b}{b + d}$$

and

$$\left(\frac{c}{a + c}\right)^\lambda \left(\frac{d}{b + d}\right)^{1-\lambda} \leq \lambda \frac{c}{a + c} + (1 - \lambda) \frac{d}{b + d}.$$

Hence,

$$\left(\frac{a}{a + c}\right)^\lambda \left(\frac{b}{b + d}\right)^{1-\lambda} + \left(\frac{c}{a + c}\right)^\lambda \left(\frac{d}{b + d}\right)^{1-\lambda} \leq \lambda + (1 - \lambda) = 1,$$

which proves the statement in the lemma.  $\square$

One of the important definitions which will be frequently used throughout the paper is that of a convex transformation. We say that  $f_1$  is a (strictly) convex transformation of  $f_2$  if there exists a (strictly) convex function  $f$  such that  $f_1(x) = (f \circ f_2)(x) = f(f_2(x))$ .

**Lemma 3.** Let  $f_1, f_2: I \rightarrow \mathbb{R}$  such that  $f_2^{-1}$  exists. Then  $f_1$  is a (strictly) convex transformation of  $f_2$  if and only if the composition  $f_1 \circ f_2^{-1}$  is (strictly) convex.

*Proof.* See [18].  $\square$

Recall also that a function  $f: I \rightarrow \mathbb{R}$  is called *concave* if and only if  $-f$  is convex. Thus a function  $f: I \rightarrow \mathbb{R}$  is *log-concave* if and only if  $1/f$  is log-convex. Therefore, the definitions and results stated in this section can be easily adapted for (log-)concavity.

## 2.2 Preferences

Let  $X \subset \mathbb{R}_+$  be the set of outcomes. We will assume that  $X$  is an interval of non-negative real numbers containing 0. The natural interpretation is that outcomes are monetary (for an infinitely divisible currency) but this is not essential. Let  $T = [0, \infty)$  be a set of points in time where 0 corresponds to the present moment. The Cartesian product  $X \times T$  will be identified with the set of timed outcomes, i.e., a pair  $(x, t) \in X \times T$  is understood as a dated outcome, when a decision-maker receives  $x$  at time  $t$  and nothing at all other time periods in  $T \setminus t$ .

Suppose that a decision-maker has a preference order  $\succsim$  on the set of timed outcomes with  $\succ$  expressing strict preference and  $\sim$  indifference. We say that a utility function  $U: X \times T \rightarrow \mathbb{R}$  *represents* the preference order  $\succsim$ , if for all  $x, y \in X$  and all  $t, s \in T$  we have  $(x, t) \succsim (y, s)$  if and only if  $U(x, t) \geq U(y, s)$ . This is a *discounted utility (DU) representation* if

$$U(x, t) = D(t)u(x), \quad (5)$$

where  $u: X \rightarrow \mathbb{R}$  is a continuous and strictly increasing function with  $u(0) = 0$ , and  $D: T \rightarrow (0, 1]$  is continuous and strictly decreasing such that  $D(0) = 1$  and  $\lim_{t \rightarrow \infty} D(t) = 0$ .

The function  $u$  is called the *instantaneous utility function*, and  $D$  is called the *discount function* associated with  $\succsim$ . We say that the pair  $(u, D)$  provides a *discounted utility representation* for  $\succsim$ . Fishburn and Rubinstein [7] provide an axiomatic foundation for a discounted utility representation. A list of their axioms is given in the Appendix. We assume that  $\succsim$  has a discounted utility representation throughout the paper.

As  $D$  is strictly decreasing, our decision maker is always impatient. However, as time goes by, her impatience may increase or decrease.

**Definition 1** ([19]). *The preference order  $\succsim$  exhibits (strictly) decreasing impatience (DI) if for all  $\sigma > 0$ , all  $0 \leq t < s$  and all outcomes  $y > x > 0$ , the equivalence  $(x, t) \sim (y, s)$  implies  $(x, t + \sigma) \precsim (y, s + \sigma)$ .*

Increasing impatience (II) can be defined by reversing the final preference ranking in Definition 1. However, we focus on DI preferences in the present paper, since this appears to be the empirically relevant case. As in the previous sentence, we also use the acronym “DI” interchangeably as a noun (“decreasing impatience”) and an adjective (“decreasingly impatient”), relying on context to indicate the intended meaning.

In case the preference order  $\succsim$  has a discounted utility representation, the characterization of DI in terms of the discount function is well-known.<sup>1</sup>

**Proposition 4** ([10, 19]). *Let  $\succsim$  be a preference order having discounted utility representation with the discount function  $D$ . The following conditions are equivalent:*

- *The preference order  $\succsim$  exhibits (strictly) DI ;*
- *$D$  is (strictly) log-convex on  $[0, \infty)$ .*

---

<sup>1</sup>The proof in [10, Theorem 3.3] can be easily adapted to demonstrate an analogous result for increasing impatience: the preference order  $\succsim$  exhibits (strictly) II if and only if  $D$  is (strictly) log-concave on  $[0, \infty)$ .

We say that discount function  $D$  is (strictly) DI if the preference order  $\succsim$  exhibits (strictly) DI and has a discounted utility representation with discount function  $D$ .

We next show that a preference order  $\succsim$  with a discounted utility representation exhibits strictly DI if and only if it exhibits present bias in the sense of Jackson and Yariv [11, p. 4190]. It is important to note, however, that Jackson and Yariv assume a discrete time setting, whereas we allow time to be continuous. The following Definition 2 is, therefore, the continuous-time analogue of their present bias definition.<sup>2</sup>

**Definition 2** (Present Bias). *The preference order is present-biased if*

- (i)  $(x, t) \preceq (y, s)$  implies  $(x, t + \sigma) \preceq (y, s + \sigma)$  for every  $x, y$ , every  $\sigma > 0$  and every  $s, t \in T$  such that  $s > t \geq 0$ ; and
- (ii) for every  $s, t \in T$  with  $s > t \geq 0$  and every  $\sigma > 0$  there exist  $x^*$  and  $y^*$  such that  $(x^*, t + \sigma) \prec (y^*, s + \sigma)$  and  $(x^*, t) \succ (y^*, s)$ .

Proposition 5 gives conditions which are equivalent to present bias for preferences with a discounted utility representation:

**Proposition 5.** *Suppose that  $\succsim$  has a discounted utility representation. Then the first condition of Definition 2 is equivalent to convexity of  $\ln D(t)$ ; while the second condition of Definition 2 is equivalent to strict convexity of  $\ln D(t)$ .*

*Proof.* We start by proving the first equivalence. Since a discounted utility representation exists, the first condition is equivalent to:

$$u(x)D(t) \leq u(y)D(s) \quad \text{implies} \quad u(x)D(t + \sigma) \leq u(y)D(s + \sigma)$$

for every  $x, y$ , every  $\sigma > 0$  and every  $s, t \in T$  with  $s > t \geq 0$ . This may be rewritten as follows:

$$u(x) \leq \frac{D(s)}{D(t)}u(y) \quad \text{implies} \quad u(x) \leq \frac{D(s + \sigma)}{D(t + \sigma)}u(y).$$

Since  $(y, s) \succ (x, t)$ ,  $s > t$  and  $D$  is strictly decreasing it follows that  $u(y) > u(x)$ . As  $u(0) = 0$  and  $u$  is strictly increasing we deduce that  $u(y) > 0$ . Since  $u$  is continuous,  $x$  and  $y$  can be chosen so that  $u(x)/u(y)$  takes any value in  $[0, 1)$ . We therefore have:

$$\frac{D(s)}{D(t)} \leq \frac{D(s + \sigma)}{D(t + \sigma)}.$$

Alternatively,

$$\ln D(s) + \ln D(t + \sigma) \leq \ln D(s + \sigma) + \ln D(t) \tag{6}$$

for every  $\sigma > 0$  and  $s, t \in T$  with  $s > t \geq 0$ . Inequality (6) is equivalent to convexity of  $\ln D(t)$ .

---

<sup>2</sup>There is an inconsistency between Jackson and Yariv's Present Bias definition in their 2014 paper (referenced here) and their 2015 paper [12]. We adhere to the former definition.

The second part is proved analogously. Under a discounted utility representation the second condition is equivalent to the following: for every  $s, t \in T$  with  $s > t \geq 0$  and every  $\sigma > 0$  there exist  $x^*$  and  $y^*$  such that:

$$u(x^*)D(t + \sigma) < u(y^*)D(s + \sigma) \quad \text{but} \quad u(x^*)D(t) > u(y^*)D(s).$$

Equivalently,

$$\frac{D(s)}{D(t)}u(y^*) < u(x^*) < \frac{D(s + \sigma)}{D(t + \sigma)}u(y^*).$$

From the fact that  $(y^*, s + \sigma)$  is preferred to  $(x^*, t + \sigma)$  with  $s > t$  we deduce that  $u(y^*) > 0$ . Hence,

$$\frac{D(s)}{D(t)} < \frac{D(s + \sigma)}{D(t + \sigma)}.$$

This inequality is equivalent to:

$$\ln D(s) + \ln D(t + \sigma) < \ln D(s + \sigma) + \ln D(t) \tag{7}$$

for every  $s, t \in T$  with  $s > t \geq 0$  and every  $\sigma > 0$ . Inequality (7) holds if and only if  $\ln D(t)$  is *strictly* convex.  $\square$

Proposition 5 implies that when a discounted utility representation exists the first condition of Definition 2 follows from the second one, since *strict* convexity of  $\ln D(t)$  implies convexity of  $\ln D(t)$ . An immediate consequence is that present bias is equivalent to strictly DI, as stated below:

**Corollary 6.** *Suppose the preference order  $\succsim$  admits a discounted utility representation. Then it exhibits present bias if and only if  $\succsim$  exhibits strictly DI.*

## 3 Comparative DI

### 3.1 More DI and log-convexity

Assume now that there are two decision makers and they are both decreasingly impatient. What does it mean to say that one of them is more decreasingly impatient than the other? The answer to this question is in the following definition:<sup>3</sup>

**Definition 3** (cf. [19], Definition 2; [4], Definition 1). *We say that  $\succsim_1$  exhibits [strictly] more DI than  $\succsim_2$ , if for every  $\sigma > 0$ , every  $\rho$ , every  $s, t \in T$  with  $0 \leq t < s$  and every  $x, x', y, y' \in X$  with  $y > x > 0$  and  $y' > x' > 0$ , the conditions  $(x', t) \sim_2 (y', s)$ ,  $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$  and  $(x, t) \sim_1 (y, s)$  imply  $(x, t + \sigma) \precsim_1 [\prec_1] (y, s + \sigma + \rho)$ .*

Not surprisingly, the (strictly) more DI relation may be expressed in terms of the comparative convexity of the logarithms of the respective discount functions, for cases in which both preference relations have discounted utility representations.

---

<sup>3</sup>Since the sign of  $\rho$  is not restricted in Definition 3, it actually applies to preferences that exhibit decreasing or increasing impatience.



**Proposition 7** (cf. [19], Proposition 1). *Let  $\succsim_1$  and  $\succsim_2$  be two preference orders with discounted utility representation by  $(u_1, D_1)$  and  $(u_2, D_2)$ , respectively. The following conditions are equivalent:*

- (i) *The preference order  $\succsim_1$  exhibits (strictly) more DI than  $\succsim_2$ ;*
- (ii)  *$\ln D_1(D_2^{-1}(e^z))$  is (strictly) convex in  $z$  on  $(-\infty, 0]$ .*

*Proof.* See the Appendix. We follow Prelec’s argument for his Proposition 1 in [19]. The additional adjustment is the necessity to replace convexity of the log-transformed discount function with strict convexity for the strictly more DI case. The required adjustments are not substantial but we have included a detailed proof as it clarifies some details omitted from Prelec’s original version [19].  $\square$

Note that the form of the utility functions  $u_1$  and  $u_2$  does not influence the comparative DI properties of preference relations.

**Corollary 8.** *Let  $\succsim_1$  and  $\succsim_2$  be two preference relations with discounted utility representations  $(u_1, D_1)$  and  $(u_2, D_2)$ , respectively, where  $D_2(t) = \delta^t$  and  $\delta \in (0, 1)$ . The preference order  $\succsim_1$  exhibits (strictly) DI if and only if it exhibits (strictly) more DI than  $\succsim_2$ .*

*Proof.* Prelec [19] proves that a preference relation is DI if and only if it is more DI than an exponential discount function. We prove the “strict” part of the claim.

Since  $D_2(t) = \delta^t$  and  $\delta \in (0, 1)$  we have

$$D_2^{-1}(e^z) = \frac{z}{\ln \delta} \geq 0.$$

By Proposition 7, for  $\succsim_1$  to exhibit strictly more DI than  $\succsim_2$  it is necessary and sufficient that  $\ln D_1(D_2^{-1}(e^z))$  is strictly convex in  $z$  on  $(-\infty, 0]$ . However,

$$\ln D_1(D_2^{-1}(e^z)) = \ln D_1\left(\frac{z}{\ln \delta}\right) = \ln D_1(t),$$

where

$$t = \frac{z}{\ln \delta} \in [0, \infty)$$

when  $z$  takes arbitrary values in  $(-\infty, 0]$ . Therefore, strict convexity of  $\ln D_1(D_2^{-1}(e^z))$  in  $z$  on  $(-\infty, 0]$  is equivalent to strict convexity of  $\ln D_1(t)$  in  $t$  on  $[0, \infty)$ . By Proposition 4, strict convexity of  $\ln D_1(t)$  in  $t$  on  $[0, \infty)$  is equivalent to  $\succsim_1$  exhibiting strictly DI.  $\square$

The following notations will be used below:

- If  $D_1$  and  $D_2$  represent equally DI preferences, we write  $D_1 \sim_{DI} D_2$ ;
- If  $D_1$  represents more DI preferences than  $D_2$ , we write  $D_1 \succsim_{DI} D_2$ ;
- If  $D_1$  represents strictly more DI preferences than  $D_2$ , we write  $D_1 \succ_{DI} D_2$ .

The following corollary, due to Prelec [19], characterizes the relation between any two discount functions from the same DI class.

**Corollary 9** ([19]). *For any two discount functions  $D_1$  and  $D_2$ , we have  $D_1 \sim_{DI} D_2$  if and only if  $D_1(t) = D_2(t)^c$ , where  $c > 0$  is a constant not depending on  $t$ .*

The  $\succsim_{DI}$  relation is a partial order. In fact, the “more DI” and “strictly more DI” relations are both transitive. This is established in the following proposition.

**Proposition 10.** *If  $D_1 \succsim_{DI} D_2$  and  $D_2 \succsim_{DI} D_3$ , then  $D_1 \succsim_{DI} D_3$ . If at least one of the relations  $D_1 \succsim_{DI} D_2$  or  $D_2 \succsim_{DI} D_3$  is strict, then  $D_1 \succ_{DI} D_3$ .*

*Proof.* Suppose  $D_1 \succsim_{DI} D_2$  and  $D_2 \succsim_{DI} D_3$ . By Proposition 7, we know that both  $\ln D_1(D_2^{-1}(e^z))$  and  $\ln D_2(D_3^{-1}(e^z))$  are convex in  $z$  on  $(-\infty, 0]$ . Defining  $h_i = \ln D_i$  for  $i \in \{1, 2, 3\}$ , we can equivalently state that  $h_1 \circ h_2^{-1}$  and  $h_2 \circ h_3^{-1}$  are convex on  $(-\infty, 0]$ . To prove transitivity it is sufficient to show that  $\ln D_1(D_3^{-1}(e^z))$  is convex in  $z$  on  $(-\infty, 0]$ , or equivalently that  $h_1 \circ h_3^{-1}$  is convex on  $(-\infty, 0]$ .

Let  $f_1 = h_1 \circ h_2^{-1}$  and  $f_2 = h_2 \circ h_3^{-1}$ . Then

$$\ln D_1(D_3^{-1}(e^z)) = h_1(h_3^{-1}(z)) = h_1 h_2^{-1}(h_2 h_3^{-1}(z)) = f_1 \circ f_2(z) = f(z).$$

By the assumption,  $f_1$  and  $f_2$  are convex functions. Note that  $f_1$  is increasing, as the composition of two decreasing functions  $h_1$  and  $h_2^{-1}$ . Indeed,  $h_1 = \ln D_1$  is a strictly decreasing function as  $D_1$  is strictly decreasing, and  $h_2^{-1}$  is a decreasing function as the inverse of the decreasing function  $h_2$ . Lemma 1 then implies that  $f(z) = f_1 \circ f_2(z) = \ln D_1(D_3^{-1}(e^z))$  is convex, and that  $f$  is strictly convex if  $f_i$  is strictly convex for some  $i \in \{1, 2\}$ .  $\square$

### 3.2 Time Preference Rate and Index of DI

In this section we assume that  $D$  is twice continuously differentiable. The *rate of time preference*,  $r(t)$ , is defined as follows:

$$r(t) = -\frac{D'(t)}{D(t)}.$$

The following lemma relates the DI property to the behaviour of  $r(t)$ .<sup>4</sup>

**Lemma 11.** *Let  $\succsim$  be a preference relation with DU representation  $(u, D)$  in which  $D$  is twice continuously differentiable. Then the following conditions are equivalent:*

- (i) *The preference relation exhibits (strictly) DI;*

---

<sup>4</sup>Takeuchi [26] contains a related result. His Corollary 1 says that the *hazard function* is decreasing (increasing) if and only if preferences exhibit decreasing (increasing) impatience. Takeuchi’s hazard function  $h(t)$  corresponds to our time preference rate  $r(t)$ . However, Takeuchi does not analyse the case of strictly decreasing impatience.

(ii) The time preference rate  $r(t)$  is (strictly) decreasing on  $[0, \infty)$ .

*Proof.* Suppose that  $r(t)$  is decreasing on  $[0, \infty)$ . This is equivalent to

$$r'(t) = -\frac{D''(t)D(t) - (D'(t))^2}{D(t)^2} = \frac{(D'(t))^2 - D''(t)D(t)}{D(t)^2} \leq 0.$$

Or, alternatively,  $D''D - (D')^2 \geq 0$ . This inequality is equivalent to log-convexity of  $D$ , which, by Proposition 4, means that the preference order exhibits DI.

To prove the equivalence of strictly DI preferences and a strictly decreasing rate of time preference, recall that a continuously differentiable function  $r: \mathbb{R}_+ \rightarrow \mathbb{R}$  is strictly decreasing if and only if  $r'(t) \leq 0$  for all  $t$  and the set  $\{t \mid r'(t) = 0\}$  contains no non-trivial interval [25, 23]. If a function  $v$  is differentiable on an open interval  $I \subset \mathbb{R}$ , then  $v$  is strictly convex on  $I$  if and only if  $v'$  is strictly increasing on  $I$  [23]. Assume that  $r(t)$  is strictly decreasing on  $[0, \infty)$ . Let  $M \subseteq \mathbb{R}_+$  be the set of  $t$  values such that  $r'(t) < 0$ . Then  $D''(t)D(t) - [D'(t)]^2 > 0$  for all  $t \in M$ . Since  $\mathbb{R}_+ \setminus M$  contains no non-trivial interval,  $r'(t)$  being strictly decreasing is equivalent to  $D$  being strictly log-convex.  $\square$

One way to measure the level of DI for suitably differentiable discount functions was suggested by Prelec [19]. Since more DI preferences have discount functions which are more log-convex, the natural criterion would be some measure of convexity of the log of the discount function. The Arrow-Pratt coefficient, which is a measure of the concavity of a function, can be adapted to this purpose. Indeed, a non-increasing rate of time preference,  $r'(t) \leq 0$ , is precisely analogous to the notion of decreasing risk aversion in Pratt [18].

Recall that  $D$  is a twice continuously differentiable function. The associated *rate of impatience*,  $IR(D)$ , is defined as follows:

$$IR(D) = -\frac{D''}{D'}.$$

The *index of DI* of  $D$ , denoted  $I_{DI}(D)$ , is the difference between the rate of impatience and the rate of time preference:

$$I_{DI}(D) = IR(D) - r(D) = \left(-\frac{D''}{D'}\right) - \left(-\frac{D'}{D}\right).$$

Note that

$$I_{DI}(D)(t) = -\frac{r'(t)}{r(t)} = -\frac{d}{dt} \ln[r(t)]. \quad (8)$$

Prelec [19] proved that if  $\succsim_1$  and  $\succsim_2$  both have DU representations with twice continuously differentiable discount functions,  $D_1$  and  $D_2$  respectively, then  $\succsim_1$  exhibits more DI than  $\succsim_2$  if and only if  $I_{DI}(D_1) \geq I_{DI}(D_2)$  on the interval  $[0, \infty)$ . The following proposition strengthens this result.

**Proposition 12.** *Let  $\succsim_1$  and  $\succsim_2$  have DU representations with discount functions  $D_1$  and  $D_2$ , respectively, where  $D_1$  and  $D_2$  are twice continuously differentiable. Then the preference order  $\succsim_1$  exhibits strictly more DI than  $\succsim_2$  if and only if  $I_{DI}(D_1) \geq I_{DI}(D_2)$  on the interval  $[0, \infty)$  and  $\{t \mid I_{DI}(D_1)(t) = I_{DI}(D_2)(t)\}$  contains no non-trivial interval.*

*Proof.* Prelec's [19, Proposition 2] proof applies the Arrow-Pratt coefficient [18], which is used to compare the concavity of functions. There is no straightforward adaptation of Prelec's argument to the case of *strict* concavity. We therefore adapt Pratt's [18] original argument directly.

Recall that  $D_1$  is strictly more DI than  $D_2$  if and only if  $\ln(D_1)$  is strictly more convex than  $\ln(D_2)$  on  $[0, \infty)$ . Let  $h_1 = \ln(D_1)$  and  $h_2 = \ln(D_2)$ , so  $h_1$  and  $h_2$  are strictly decreasing functions. The function  $h_1$  is strictly more convex than  $h_2$  on  $(-\infty, 0]$  if and only if there exists a strictly convex transformation  $f$  such that  $h_1 = f(h_2)$ , or, equivalently,  $h_1(h_2^{-1}(z))$  is strictly convex on  $(-\infty, 0]$ .

The first derivative of  $h_1(h_2^{-1}(z))$  is:

$$\frac{dh_1(h_2^{-1}(z))}{dz} = \frac{h_1'(h_2^{-1}(z))}{h_2'(h_2^{-1}(z))}. \quad (9)$$

We need to show that expression (9) is strictly increasing. Note that  $h_2^{-1}(z)$  is strictly decreasing since  $h_2$  is strictly decreasing. Therefore, (9) is strictly increasing if and only if  $h_1'(x)/h_2'(x)$  is strictly decreasing. The latter is satisfied if and only if

$$\log \left[ \frac{h_1'(x)}{h_2'(x)} \right] \quad (10)$$

strictly decreases (since  $\log(x)$  is strictly increasing). The first derivative of (10) is:

$$\frac{h_2'(x)}{h_1'(x)} \cdot \frac{h_1''(x)h_2'(x) - h_1'(x)h_2''(x)}{[h_2'(x)]^2} = \frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)}$$

Therefore (10) is strictly decreasing if and only if

$$\frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} \leq 0$$

and the set

$$\left\{ x \mid \frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} = 0 \right\}$$

contains no non-trivial interval.

Note that:

$$\frac{h_i''}{h_i'} = \frac{D_i''}{D_i'} - \frac{D_i'}{D_i}.$$

Therefore,

$$\frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} \leq 0$$

is equivalent to

$$-\frac{D_1''}{D_1'} - \left(-\frac{D_1'}{D_1}\right) \geq -\frac{D_2''}{D_2'} - \left(-\frac{D_2'}{D_2}\right).$$

This means that  $D_1 \succ_{DI} D_2$  if and only if  $I_{DI}(D_1) \geq I_{DI}(D_2)$  on  $[0, \infty)$ , and  $\{t \mid I_{DI}(D_1)(t) = I_{DI}(D_2)(t)\}$  contains no non-trivial interval.  $\square$

From Proposition 12, Lemma 11 and (8) it follows that  $\succ$  is DI if and only if  $I_{DI}(D) \geq 0$  on  $[0, \infty)$ , and  $\succ$  is strictly DI if and only if  $I_{DI}(D) \geq 0$  on  $[0, \infty)$  and  $\{t \mid I_{DI}(D)(t) = 0\}$  contains no non-trivial interval. Note that the index of DI equals zero for an exponential discount function.<sup>5</sup>

The following example illustrates the index of DI for a generalized hyperbolic discount function. We will make use of this information later.

**Example 1.** *The function  $D(t) = (1 + ht)^{-\alpha/h}$ , with  $h > 0$  and  $\alpha > 0$ , is called the generalized hyperbolic discount function. For this function we have:*

$$r(t) = \alpha(1 + ht)^{-1} \quad \text{and} \quad IR(D)(t) = (\alpha + h)(1 + ht)^{-1}.$$

Therefore,  $I_{DI}(D)(t) = h(1 + ht)^{-1}$ . If  $D_1(t) = (1 + h_1t)^{-\alpha/h_1}$  and  $D_2(t) = (1 + h_2t)^{-\alpha/h_2}$  are two generalized hyperbolic discount functions then  $D_1 \succ_{DI} [ \succ_{DI} ] D_2$ , if and only if  $h_1 \geq [ > ] h_2$ .

Thus the parameter  $h$  may be used as a measure of the degree of DI of a generalized hyperbolic discount function, while the parameter  $\alpha$  has no influence on  $I_{DI}(D)$ . We call parameter  $h$  the hyperbolic discount rate. The special case of a generalized hyperbolic discount function with  $\alpha = h > 0$  is called the proportional hyperbolic discount function.

## 4 Mixtures of Discount Functions

As described in the introduction, there are some situations in which the necessity arises to calculate a convex combination (mixture) of discount functions.

### 4.1 Two scenarios

The first situation where a convex combination of discount functions may be used is when there is a group of decision makers with different discount functions and a social discount function needs to be constructed. The natural option is averaging the discount functions among individuals, which is equivalent to averaging the discounted utilities when all agents have identical utility functions. This approach has been widely used in the existing literature on time preferences ([11]).

Indeed, suppose that we have a set of agents  $M = \{1, \dots, m\}$ . Assume that agent  $i$  has time preferences with DU representation  $(u, D_i)$ . Thus, all agents have the same

---

<sup>5</sup>Similarly,  $\succ$  is II if and only if  $I_{DI}(D) \leq 0$  on  $[0, \infty)$  and strictly II if and only if  $I_{DI}(D) \leq 0$  on  $[0, \infty)$  and  $\{t \mid I_{DI}(D)(t) = 0\}$  contains no non-trivial interval.

instantaneous utility function. Then we define the *collective (utilitarian) utility function* as  $\hat{u}(x) = mu(x)$  and the *collective total utility* at time  $t$  is

$$\hat{U}(x, t) = \sum_{i=1}^m D_i(t)u(x) = \left( \frac{1}{m} \sum_{i=1}^m D_i(t) \right) \hat{u}(x).$$

Thus, we obtain the *collective discount function*  $D = \frac{1}{m} \sum_{i=1}^m D_i$ .

In the second possible scenario, discussed by Sozou [24] and Weitzman [27], there is a single decision maker with some uncertainty about her discount function. For example, there may be some possibility of not surviving to any given period,  $t$ , described by a survival function with and uncertain (constant) hazard rate [24]. Then the expected discount function of this decision maker can be calculated as a weighted average of the distinct discount functions that may eventuate.

If the discount function  $D_i$  eventuates with probability  $p_i$ , then the expected utility of the decision maker is

$$\hat{U}(x, t) = \sum_{i=1}^m p_i D_i(t)u(x) = \left( \sum_{i=1}^m p_i D_i(t) \right) u(x).$$

and the *certainty equivalent discount function* will be  $D = \sum_{i=1}^m p_i D_i$ .

The same question arises in both cases: Is it possible to make some conclusion about the behaviour of the convex combination of distinct discount functions in comparison with its components, if all the component discount functions exhibit DI?

## 4.2 Mixtures of discount functions with decreasing impatience

Given a set of discount functions  $\{D_1, D_2, \dots, D_n\}$ , we define a *mixture* of them as

$$D = \sum_{i=1}^n \lambda_i D_i,$$

where  $0 < \lambda_i < 1$  for all  $i$  and  $\sum_{i=1}^n \lambda_i = 1$ . Note that we define a mixture such that each  $D_i$  has a *strictly* positive weight.

We first discuss some known results related to the mixture of discount functions.

One of the most recent results was obtained by Jackson and Yariv [11], who demonstrated that if all decision makers in a group have exponential discount functions, but they do not all have the same discount factor, then their collective discount function must be present biased.

It has also been noted by several authors (for example, [20] and [22]), that time preferences have strong similarities with risk preferences and that some results from risk theory are relevant in the context of intertemporal choice. Pratt [18] showed that decreasing risk aversion is preserved under linear combinations. As was observed in Section 3.2, decreasing risk aversion is analogous to non-decreasing time preference rate, or DI of the discount function. Therefore, Pratt's result can be translated into our time preference framework as follows:

**Proposition 13** (cf. [18], Theorem 5). *Let  $\succsim_1, \succsim_2, \dots, \succsim_n$  have DU representations with twice continuously differentiable discount functions  $D_1, \dots, D_n$ , respectively. Assume that  $\succsim_1, \succsim_2, \dots, \succsim_n$  all exhibit DI. Let*

$$D = \sum_{i=1}^n \lambda_i D_i,$$

*be a mixture of  $D_1, \dots, D_n$ . Then  $D$  is DI. It is strictly DI if and only if*

$$\{ t \mid r_1(t) = r_2(t) = \dots = r_n(t) \text{ and } r'_1(t) = r'_2(t) = \dots = r'_n(t) = 0 \}$$

*contains no non-trivial interval.*

*Proof.* From the definition of time preference rate it follows that  $D'_i = -r_i D_i$  for all  $i = 1, \dots, n$ . The time preference rate for  $D$  is:

$$r = -\frac{D'}{D} = -\frac{\sum_{i=1}^n \lambda_i D'_i}{D} = \sum_{i=1}^n \frac{\lambda_i D_i}{D} r_i.$$

By Lemma 11, to prove that  $D$  exhibits DI we must show that  $r'(t) \leq 0$ :

$$r' = \sum_{j=1}^n \frac{\lambda_j D'_j \sum_{i=1}^n \lambda_i D_i - \lambda_j D_j \sum_{i=1}^n \lambda_i D'_i}{D^2} r_j + \sum_{j=1}^n \frac{\lambda_j D_j}{D} r'_j.$$

Rearranging and substituting  $D'_i = -r_i D_i$  we obtain:

$$r' = \sum_{j=1}^n \frac{\lambda_j D_j}{D} r'_j + \frac{Q}{D^2},$$

where

$$Q = - \sum_{j=1}^n \left[ \lambda_j r_j D_j \sum_{i=1}^n \lambda_i D_i - \lambda_j D_j \sum_{i=1}^n \lambda_i r_i D_i \right] r_j$$

This is a quadratic form in  $D_1, D_2, \dots, D_n$  with the coefficient on  $D_i D_j$  being

$$\lambda_i \lambda_j (r_i r_j - r_j^2) + \lambda_i \lambda_j (r_i r_j - r_i^2) = -\lambda_i \lambda_j (r_i - r_j)^2.$$

Hence

$$Q = - \sum_{i < j} \lambda_i \lambda_j D_i D_j (r_i - r_j)^2.$$

Since  $D_i \in (0, 1]$ ,  $\lambda_i > 0$  and  $r'_i \leq 0$  for all  $i = 1, \dots, n$ , we have  $r'(t) \leq 0$ . Therefore,  $\succsim$  is DI. The preference relation  $\succsim$  is strictly DI if and only if  $r'(t)$  is strictly decreasing. We see that  $r'(t)$  is strictly decreasing iff

$$\{ t \mid r_1(t) = r_2(t) = \dots = r_n(t) \text{ and } r'_1(t) = r'_2(t) = \dots = r'_n(t) = 0 \}$$

contains no non-trivial interval. □

The following corollary describes an important special case of Proposition 13:

**Corollary 14.** *Mixtures of non-identical exponential discount functions are strictly DI.*

Corollary 14 is therefore a continuous-time version of Jackson and Yariv (2014, Proposition 1). Prelec [19, Corollary 4] considers the mixture of two discount functions only, but does not require differentiability. He proves that the mixture of two equally DI discount functions is more DI than its components. Prelec [19, Corollary 4] implies the special case of Jackson and Yariv's [11] result when  $n = 2$ .

Our objective is to establish such a result which is more general than both Prelec [19] and Jackson and Yariv [11]. The result we obtain is stated in the following theorem:

**Theorem 1.** *Let  $n \geq 2$  and  $D_1, \dots, D_n$  be distinct discount functions such that*

$$D_1 \succ_{DI} D_2 \succ_{DI} \dots \succ_{DI} D_n.$$

*If  $D$  is a mixture of  $D_1, \dots, D_n$ , then  $D \succ_{DI} D_n$ .*

To construct the proof of this theorem, two preliminary results will be useful. The first is a strengthening of a result in Prelec [19].

**Proposition 15.** *Let  $\lambda \in (0, 1)$ . If two distinct discount functions  $D_1$  and  $D_2$  satisfy  $D_1 \sim_{DI} D_2$ , then their mixture,  $D = \lambda D_1 + (1 - \lambda)D_2$ , represents strictly more DI preferences than each  $D_i$ . That is,  $D \succ_{DI} D_1$  and  $D \succ_{DI} D_2$ .*

*Proof.* As  $D_1 \sim_{DI} D_2$ , then, by Corollary 9,  $D_1(t) = D_2(t)^c$ , where  $c \neq 1, c > 0$ . By Proposition 7, it is necessary to demonstrate strict convexity of  $f(z) = \ln D(D_1^{-1}(e^z))$  for  $z \leq 0$ . By Proposition 10 it is also sufficient. We note that:

$$\begin{aligned} f(z) &= \ln D(D_1^{-1}(e^z)) = \ln(\lambda D_1(D_1^{-1}(e^z)) + (1 - \lambda)D_2(D_1^{-1}(e^z))) \\ &= \ln(\lambda e^z + (1 - \lambda)e^{z/c}). \end{aligned}$$

The first-order derivative of  $f(z)$  is:

$$f'(z) = \frac{\lambda e^z + \frac{1}{c}(1 - \lambda)e^{z/c}}{\lambda e^z + (1 - \lambda)e^{z/c}}.$$

The second-order derivative is:

$$f''(z) = \frac{(\lambda e^z + \frac{1}{c^2}(1 - \lambda)e^{z/c})(\lambda e^z + (1 - \lambda)e^{z/c}) - (\lambda e^z + \frac{1}{c}(1 - \lambda)e^{z/c})^2}{(\lambda e^z + (1 - \lambda)e^{z/c})^2} = \frac{p(z)}{q(z)}.$$

Both the denominator  $q(z)$  and the numerator  $p(z)$  of this fraction are strictly positive. The former is obvious. To see the latter, note that the numerator  $p(z)$  can be simplified as follows:

$$\begin{aligned} p(z) &= \left( \lambda e^z + \frac{1}{c^2}(1 - \lambda)e^{z/c} \right) (\lambda e^z + (1 - \lambda)e^{z/c}) - \left( \lambda e^z + \frac{1}{c}(1 - \lambda)e^{z/c} \right)^2 \\ &= \lambda(1 - \lambda)e^{(1+\frac{1}{c})z} - \frac{2}{c}\lambda(1 - \lambda)e^{(1+\frac{1}{c})z} + \frac{1}{c^2}\lambda(1 - \lambda)e^{(1+\frac{1}{c})z} \\ &= e^{(1+\frac{1}{c})z}\lambda(1 - \lambda) \left( 1 - \frac{1}{c} \right)^2. \end{aligned}$$

Therefore,  $f$  is a strictly convex function. □



Proposition 15 is a stronger version of Corollary 4 in [19], as we show that the mixture of two discount functions that are equally DI represents *strictly* more DI preferences, rather than just more DI preferences. It is important to note that Proposition 15 cannot be directly generalized to  $n$  discount functions by induction. To obtain the path to such generalization, we will need the following lemma.

**Lemma 16.** *Let  $\lambda \in (0, 1)$ . If two distinct discount functions  $D_1$  and  $D_2$  satisfy  $D_1 \succ_{DI} D_2$ , then their mixture  $D = \lambda D_1 + (1 - \lambda) D_2$  represents strictly more DI preferences than  $D_2$ . That is,  $D \succ_{DI} D_2$ .*

*Proof.* If  $D_1 \sim_{DI} D_2$  then the conclusion follows from Proposition 15. Suppose that  $D_1$  represents strictly more DI preferences than  $D_2$ . Then, by Proposition 7,  $\ln D_1 (D_2^{-1}(e^z))$  is strictly convex on  $(\infty, 0]$ . We need to demonstrate that  $\ln D (D_2^{-1}(e^z))$  is also strictly convex on  $(\infty, 0]$ , or, equivalently, that the following function is strictly log-convex:

$$D (D_2^{-1}(e^z)) = \lambda D_1 D_2^{-1}(e^z) + (1 - \lambda) D_2 D_2^{-1}(e^z) = \lambda D_1 D_2^{-1}(e^z) + (1 - \lambda) e^z.$$

Denote  $\lambda D_1 D_2^{-1}(e^z) = f$  and  $(1 - \lambda) e^z = g$ . Then we have:

$$D (D_2^{-1}(e^z)) = f + g,$$

where  $f$  is strictly log-convex by assumption and  $g = (1 - \lambda) e^z$  is log-convex. By Lemma 2, the sum of a strictly log-convex function and a log-convex function is strictly log-convex, hence  $D (D_2^{-1}(e^z))$  is strictly log-convex.  $\square$

We are now ready to provide the proof of Theorem 1, since Lemma 16 can be straightforwardly generalized to the case of  $n$  distinct discount functions.

*Proof of Theorem 1.* We prove this statement by induction on  $n$ . By Lemma 16 the result holds for  $n = 2$ . Suppose that the statement of the theorem is true for  $n = k$ . Let

$$D^{(k+1)} = \eta_1 D_1 + \dots + \eta_{k+1} D_{k+1}$$

where  $\sum_{i=1}^{k+1} \eta_i = 1$  and each  $\eta_i \in (0, 1)$ . Then

$$\begin{aligned} D^{(k+1)} &= \eta_1 D_1 + \dots + \eta_{k+1} D_{k+1} \\ &= (1 - \eta_{k+1}) \left( \frac{\eta_1}{1 - \eta_{k+1}} D_1 + \dots + \frac{\eta_k}{1 - \eta_{k+1}} D_k \right) + \eta_{k+1} D_{k+1}. \end{aligned}$$

Let

$$D^{(k)} = \frac{\eta_1}{1 - \eta_{k+1}} D_1 + \dots + \frac{\eta_k}{1 - \eta_{k+1}} D_k.$$

By the induction hypothesis,  $D^{(k)} \succ_{DI} D_k$ . It is also known that  $D_k \succ_{DI} D_{k+1}$ , and hence, by Proposition 10,  $D^{(k)} \succ_{DI} D_{k+1}$ . However, the mixture of these two functions is exactly

$$D^{(k+1)} = (1 - \eta_{k+1}) D^{(k)} + \eta_{k+1} D_{k+1}.$$

Then, by Proposition 16,  $D^{(k+1)} \succ_{DI} D_{k+1}$ , which completes the induction step.  $\square$

### 4.3 Mixtures of twice continuously differentiable discount functions

Note that when discount functions are suitably differentiable, Theorem 1 and Proposition 12 imply that

$$I_{DI}(D) \geq \min_i \{I_{DI}(D_i)\} \quad \text{on } [0, \infty) \quad (11)$$

and the set of  $t$  values at which equality holds does not include any non-trivial interval.

Consider the following example:

**Example 2.** Let  $D_1(t) = (1 + ht)^{-2}$  be a zero-speed hyperbolic discount function [13] and  $D_2(t) = \exp(-\alpha t^{1/2})$  be a slow Weibull discount function [13]. As shown in Example 1,  $I_{DI}(D_1)(t) = h(1 + ht)^{-1} > 0$  for all  $t$ . We also have  $I_{DI}(D_2)(t) = (2t)^{-1} > 0$  for all  $t$  since

$$r_2(t) = \frac{\alpha}{2} t^{-1/2} \quad \text{and} \quad r'_2(t) = -\frac{\alpha^2}{4} t^{-3/2}.$$

Therefore, both  $D_1$  and  $D_2$  exhibit strict DI. Assume that  $h = 0.1$ . Then

$$I_{DI}(D_1)(t) - I_{DI}(D_2)(t) = \frac{0.1}{1 + 0.1t} - 0.5 \frac{1}{t} = 0.05 \frac{t - 10}{1 + 0.1t}.$$

Obviously,  $I_{DI}(D_1)(t) \leq I_{DI}(D_2)(t)$  if and only if  $t \leq 10$  and  $I_{DI}(D_1)(t) > I_{DI}(D_2)(t)$  if and only if  $t > 10$ . It follows that  $D_1$  and  $D_2$  both are from incomparable DI classes. Since  $D_1$  and  $D_2$  both exhibit strictly DI, Proposition 13 implies that their mixture  $D$  also exhibits strictly DI.

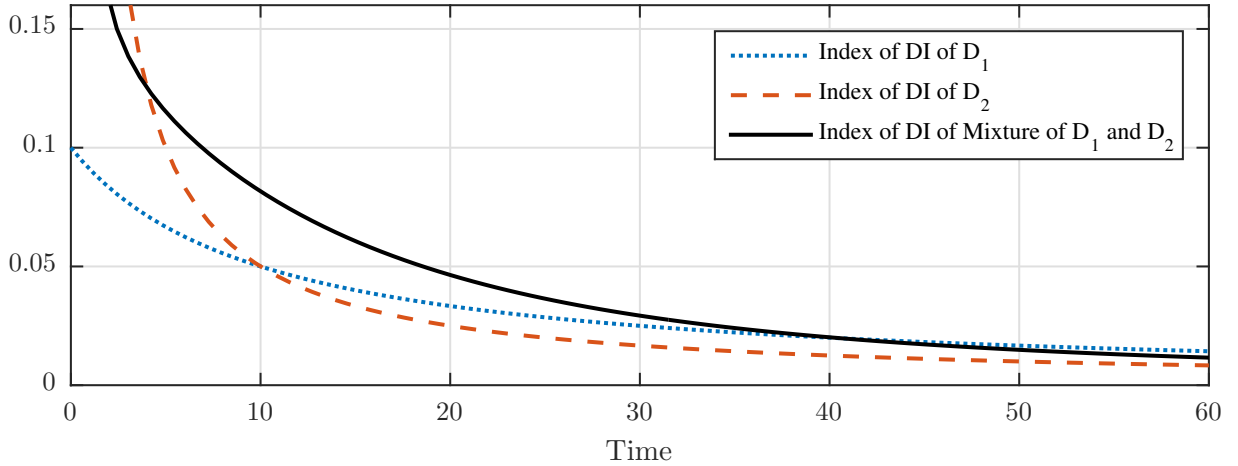


Figure 1: Index of DI for the Mixture of  $D_1$  and  $D_2$

By direct calculation we obtain the following expression:

$$I_{DI}(D)(t) = \frac{6\lambda_1 h^2(1+ht)^{-4} + 1/4\lambda_2 \alpha \exp -\alpha t^{0.5} t^{-1}(\alpha + t^{-0.5})}{2\lambda_1 h(1+ht)^{-3} + 1/2\lambda_2 \alpha \exp -\alpha t^{0.5} t^{-0.5}} \\ - \frac{2\lambda_1 h(1+ht)^{-3} + 1/2\lambda_2 \alpha \exp -\alpha t^{0.5} t^{-0.5}}{\lambda_1(1+ht)^{-2} + \lambda_2 \exp -\alpha t^{0.5}}.$$

The behaviour of  $I_{DI}(D)$  with parameters  $\lambda_1 = \lambda_2 = 0.5$ ,  $h = 0.1$  and  $\alpha = 0.12$  is illustrated in Figure 1. It can be clearly seen from Figure 1 that neither  $D \succ_{DI} D_1$  nor  $D \succ_{DI} D_2$ . However,  $I_{DI}(D) \geq \min\{I_{DI}(D_1), I_{DI}(D_2)\}$  on  $[0, \infty)$ .

Example 2 suggests that the inequality (11) may continue to hold even if the discount functions are not DI-comparable. Theorem 2 verifies this conjecture.

**Theorem 2.** Let  $\succ_1, \succ_2, \dots, \succ_n$  have DU representations with twice continuously differentiable discount functions  $D_1, D_2, \dots, D_n$ , respectively. Let  $D = \sum_{i=1}^n \lambda_i D_i$  be a mixture of  $D_1, D_2, \dots, D_n$ . Then  $I_{DI}(D) \geq \min_i \{I_{DI}(D_i)\}$  on  $[0, \infty)$ , and

$$I_{DI}(D)(t) > \min_i \{I_{DI}(D_i)(\hat{t})\}$$

if  $r_j(\hat{t}) \neq r_k(\hat{t})$  for some  $j \neq k$ .

*Proof.* Let  $I_i = I_{DI}(D_i)$  for all  $i \in \{1, \dots, n\}$  and let  $I = I_{DI}(D)$ . Recall that  $D' = -rD$  and hence  $D'' = Dr^2 - Dr' = Dr(r + I)$ . Recall also that

$$I = -\frac{D''}{D'} + \frac{D'}{D}.$$

Therefore,

$$I = \frac{-\sum_{i=1}^n \lambda_i D_i''}{\sum_{i=1}^n \lambda_i D_i'} + \frac{\sum_{i=1}^n \lambda_i D_i'}{\sum_{i=1}^n \lambda_i D_i} = \frac{\sum_{i=1}^n \lambda_i D_i r_i (r_i + I_i)}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i}.$$

This expression can be rearranged as follows:

$$I = \frac{\sum_{i=1}^n \lambda_i D_i r_i I_i}{\sum_{i=1}^n \lambda_i D_i r_i} + \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i} = \sum_{i=1}^n \alpha_i(t) I_i + Q,$$

where

$$Q = \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i} \quad \text{and} \quad \alpha_i = \frac{\lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i r_i}$$

with  $\sum_{i=1}^n \alpha_i = 1$  and  $\alpha_i \geq 0$ . Note that

$$\min_i \{I_i\} \leq \sum_{i=1}^n \alpha_i I_i \leq \max_i \{I_i\} \quad \text{for all } t \in [0, \infty).$$

The expression  $Q$  can be rewritten as:

$$Q = \frac{\left[ \sum_{i=1}^n \lambda_i D_i r_i^2 \right] \cdot \left[ \sum_{i=1}^n \lambda_i D_i \right] - \left[ \sum_{i=1}^n \lambda_i D_i r_i \right]^2}{\left[ \sum_{i=1}^n \lambda_i D_i r_i \right] \cdot \left[ \sum_{i=1}^n \lambda_i D_i \right]}.$$

The denominator of  $Q$  is strictly positive, so the sign of  $Q$  depends on the sign of the numerator. Let  $N$  be the numerator of  $Q$ :

$$N = \left[ \sum_{i=1}^n \lambda_i D_i r_i^2 \right] \cdot \left[ \sum_{i=1}^n \lambda_i D_i \right] - \left[ \sum_{i=1}^n \lambda_i D_i r_i \right]^2.$$

We can simplify  $N$  as follows:

$$N = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j D_i D_j r_i^2 - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j D_i D_j r_i r_j.$$

Therefore, we have:

$$N = \sum_{i=1}^n \sum_{j=1}^n \theta_{ij} r_i (r_i - r_j)$$

where  $\theta_{ij} = \lambda_i \lambda_j D_i D_j$ . Since  $\theta_{ij} = \theta_{ji} > 0$  for all  $i$  and  $j$  we see that

$$N = \sum_{i < j} \theta_{ij} [r_i (r_i - r_j) + r_j (r_j - r_i)] = \sum_{i < j} \theta_{ij} (r_i - r_j)^2.$$

Hence,  $N \geq 0$  and  $N > 0$  if  $r_j \neq r_k$  for some  $j \neq k$ . It follows that  $Q \geq 0$  and  $Q > 0$  if  $r_j \neq r_k$  for some  $j \neq k$ . Therefore, since  $I = \sum_{i=1}^n \alpha_i I_i + Q$  and

$$\min_i \{I_i\} \leq \sum_{i=1}^n \alpha_i I_i \leq \max_i \{I_i\} \text{ for all } t \in [0, \infty),$$

we have:

$$\min_i \{I_i\} \leq \min_i \{I_i\} + Q \leq \sum_{i=1}^n \alpha_i I_i + Q = I.$$

In other words,  $I \geq \min_i \{I_i\}$  on  $[0, \infty)$ , and  $I(\hat{t}) > \min_i \{I_i(\hat{t})\}$  if  $r_j(\hat{t}) \neq r_k(\hat{t})$  for some  $j \neq k$ .  $\square$

Observe that this result does not require discount functions to exhibit decreasing impatience. Therefore, Theorem 2 makes less restrictive assumptions than Proposition 13 – it allows the discount functions to exhibit increasing impatience.

## 4.4 Mixtures of proportional hyperbolic discount functions

Weitzman [27] shows that if different discount functions may eventuate with certain probabilities, then future costs and benefits must eventually be discounted at the lowest possible limiting time preference rate. This result is particularly salient when the possible discount functions are all exponential, with constant time preference rates. The purpose of this section is to give an analogous result for proportional hyperbolic discount functions, with constant hyperbolic discount rates (Example 1). The result in this case is very different to Weitzman's. Long-term future benefits and costs are discounted, not at the lowest hyperbolic discount rate, but at the probability-weighted harmonic mean of the individual hyperbolic discount rates.

Suppose that there is some uncertainty about the rate of time preference, and we have a set of possible scenarios  $N = \{1, \dots, n\}$  where each time preference rate  $r_i(t)$  may eventuate with probability  $p_i \geq 0$ , such that  $\sum_{i=1}^n p_i = 1$ . Since for each  $i$

$$r_i(t) = -\frac{D'_i(t)}{D_i(t)},$$

the corresponding discount function can be expressed in terms of the rate of time preference as follows

$$D_i(t) = \exp\left(-\int_0^t r_i(\tau) d\tau\right) \text{ for each } i \in N. \quad (12)$$

The certainty equivalent discount function will be:

$$D = \sum_{i=1}^n p_i D_i, \text{ where } p_i \geq 0 \text{ and } \sum_{i=1}^n p_i = 1.$$

Then the certainty equivalent time preference rate is  $r = -\frac{D'}{D}$ . Weitzman [27] proved that if each rate of time preference converges to a non-negative value as time goes to infinity, then the certainty equivalent rate of time preference converges to the lowest of these values. In other words, if  $\lim_{t \rightarrow \infty} r_i(t) = r_i^*$  with  $r_i^* \geq 0$  and  $r_1^* < r_i^*$ , where  $i \neq 1$ , then  $\lim_{t \rightarrow \infty} r(t) = r_1^*$ .

**Example 3.** Note that  $r_i(t)$  in (12) is constant if and only if  $D_i$  is exponential. In this case we have:

$$D_i(t) = \exp(-r_i t) \text{ for each } i \in N,$$

where  $r_i = \text{const.}$  Therefore, Weitzman's result implies that  $\lim_{t \rightarrow \infty} r(t) = \min_i r_i$ . Figure 2 illustrates for the case  $n = 3$ ,  $r_1 = 0.01$ ,  $r_2 = 0.02$ ,  $r_3 = 0.03$  and  $p_1 = p_2 = p_3 = 1/3$ . We also observe that the certainty equivalent rate of time preference  $r(t)$  decreases monotonically towards  $r_1$ . This is a consequence of Corollaries 8 and 14 and the fact that  $I_{DI}(D) = -r'/r$ .

However, Weitzman's result [27] does not provide much insight in the special case when each possible time preference has a DU representation with a proportional hyperbolic discount function. Suppose

$$D_i(t) = \frac{1}{1 + h_i t}$$

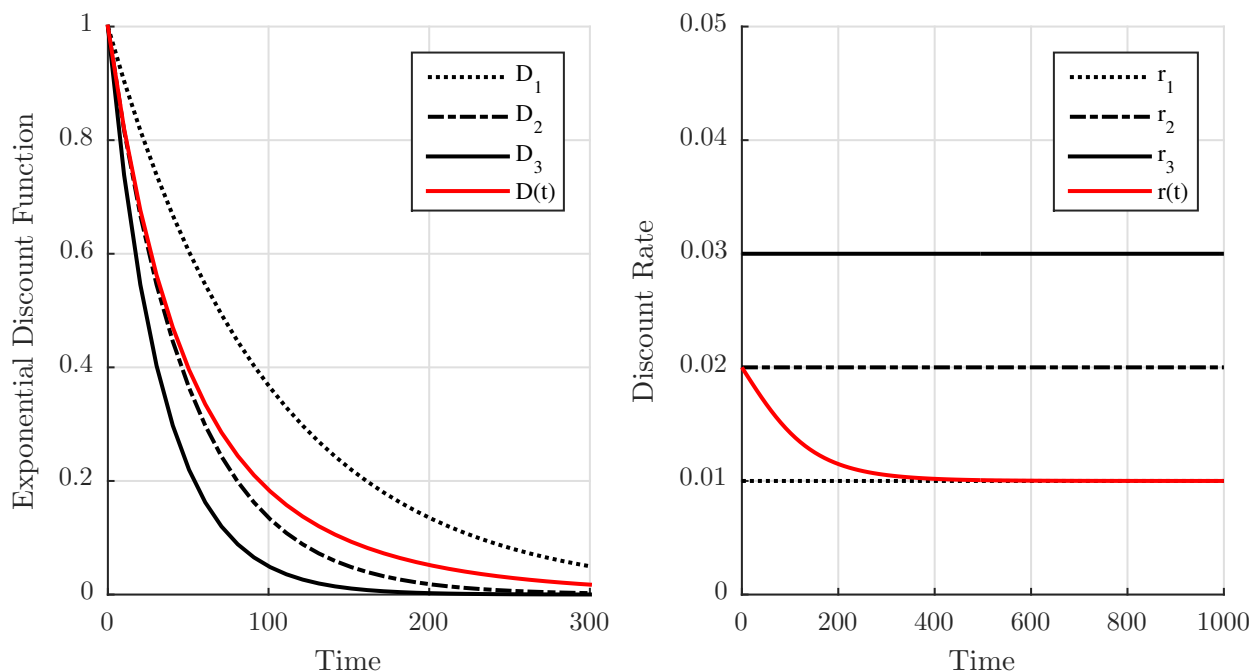


Figure 2: Mixture of Exponential Discount Functions

for each  $i \in N$ , where  $h_i > 0$  is the hyperbolic discount rate. Without loss of generality we assume that  $h_1 > h_2 > \dots > h_n$ . Suppose that  $D_i$  eventuates with probability  $p_i$  where  $p_i \geq 0$  and  $\sum_{i=1}^n p_i = 1$ . Then the certainty equivalent discount function would be

$$D(t) = \frac{p_1}{1 + h_1 t} + \dots + \frac{p_n}{1 + h_n t}.$$

The rate of time preference is

$$r_i(t) = \frac{h_i}{1 + h_i t}$$

for all  $i$ . It is obvious that  $r_i^* = r_j^* = 0$  for all  $i \neq j$  and  $\lim_{t \rightarrow \infty} r(t) = 0$ , which, indeed, corresponds to Weitzman's result. However, this conclusion does not give much information about the asymptotic behavior of the certainty equivalent discount function. Given that each possible discount function comes from a different DI class (unlike in the case of heterogeneous exponential discount functions) we would like to know which (if any) most closely characterizes the asymptotic behaviour of the certain equivalent function.

To answer this question we need to modify the analysis of Weitzman. Note that the certainty equivalent discount function can be written as

$$D(t) = \frac{1}{1 + h(t)t},$$

where  $h(t)$  is the *certainty equivalent hyperbolic discount rate*. In particular,

$$h(t) = \left( \frac{1}{D(t)} - 1 \right) \frac{1}{t},$$

so  $h(t)$  is well-defined for  $t \in (0, \infty)$ . We ask: *How does  $h(t)$  behave as  $t \rightarrow \infty$ ?*

We remind the reader that the *weighted harmonic mean* of non-negative values  $x_1, x_2, \dots, x_n$  with non-negative weights  $a_1, a_2, \dots, a_n$  satisfying  $a_1 + \dots + a_n = 1$  is

$$H(x_1, a_1; \dots; x_n, a_n) = \left( \sum_{i=1}^n \frac{a_i}{x_i} \right)^{-1}.$$

It is well-known that the weighted harmonic mean is smaller than the corresponding expected value (weighted arithmetic mean).

**Theorem 3.** *Suppose that each  $D_i$  ( $i \in N$ ) is a proportional hyperbolic discount function, with associated hyperbolic discount rate  $h_i$ . Discount function  $D_i$  will eventuate with probability  $p_i$ . Then the long-term certainty equivalent hyperbolic discount rate is the probability-weighted harmonic mean of the individual hyperbolic discount rates,  $H(h_1, p_1; \dots; h_n, p_n)$ .*

*Proof.* We note that

$$\frac{p_i}{1 + h_i t} = \frac{p_i}{h_i t} + \epsilon_i(t),$$

where  $\epsilon_i(t)/t^2 \rightarrow 0$  when  $t \rightarrow \infty$ . Let  $\epsilon(t) = \epsilon_1(t) + \dots + \epsilon_n(t)$ . Hence it follows that:

$$\begin{aligned} \frac{1}{1 + h(t)t} &= \sum_{i=1}^n p_i D_i(t) = \frac{p_1}{1 + h_1 t} + \dots + \frac{p_n}{1 + h_n t} \\ &= \frac{p_1}{h_1 t} + \dots + \frac{p_n}{h_n t} + \epsilon(t) \\ &= \left( \frac{p_1}{h_1} + \dots + \frac{p_n}{h_n} \right) \frac{1}{t} + \epsilon(t) \\ &= \frac{1}{H(h_1, p_1; \dots; h_n, p_n)t} + \epsilon(t) \\ &= \frac{1}{1 + H(h_1, p_1; \dots; h_n, p_n)t} + \hat{\epsilon}(t), \end{aligned}$$

where  $\hat{\epsilon}(t)/t^2 \rightarrow 0$  as  $t \rightarrow \infty$ . This implies that  $h(t) \rightarrow H(h_1, p_1; \dots; h_n, p_n)$  as  $t \rightarrow \infty$ .  $\square$

Figure 3 illustrates Theorem 3 for the case  $n = 3$ , when hyperbolic rates  $h_1 = 0.01$ ,  $h_2 = 0.02$  and  $h_3 = 0.03$  eventuate with equal probabilities. Note that  $h_2 = 0.02$  corresponds to the arithmetic mean of  $h_1$ ,  $h_2$  and  $h_3$ . Figure 3 displays the convergence of the certainty equivalent hyperbolic discount rate to the weighted harmonic mean  $H(h_1, p_1; h_2, p_2; h_3, p_3)$ . It also shows the certainty equivalent hyperbolic discount rate decreasing monotonically. The following proposition proves that this is always the case.

**Proposition 17.** *Suppose that each  $D_i$  ( $i \in N$ ) is a proportional hyperbolic discount function, with associated hyperbolic discount rate  $h_i$ . Discount function  $D_i$  will eventuate with probability  $p_i$ . Then the certainty equivalent hyperbolic discount rate is strictly decreasing on  $(0, \infty)$ .*

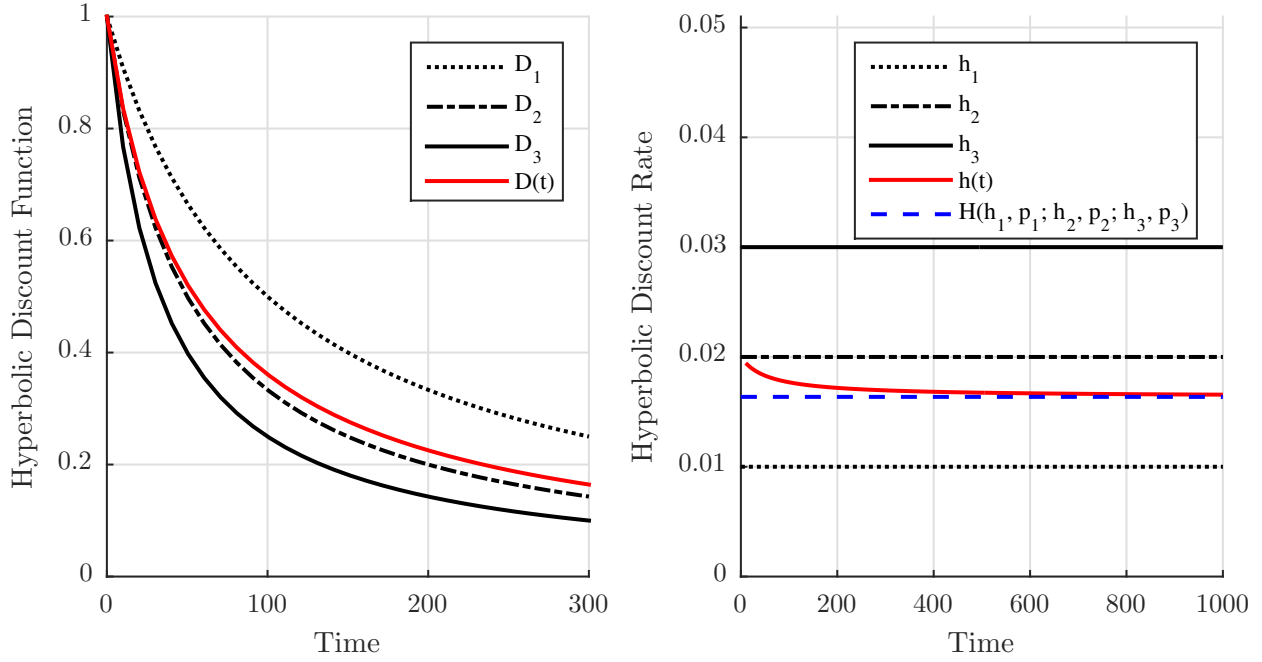


Figure 3: Mixture of Hyperbolic Discount Functions

*Proof.* We prove this statement by induction on  $n$ . First we need to prove that the statement holds for  $n = 2$ . The respective certainty equivalent hyperbolic discount rate is:

$$h(t) = \left[ \frac{1}{p_1(1 + h_1 t)^{-1} + p_2(1 + h_2 t)^{-1}} - 1 \right] \frac{1}{t}$$

for each  $t > 0$ . Rearranging:

$$h(t) = \left[ \frac{(1 + h_1 t)(1 + h_2 t)}{p_1(1 + h_2 t) + p_2(1 + h_1 t)} - 1 \right] \frac{1}{t} = \left[ \frac{1 + (h_1 + h_2)t + h_1 h_2 t^2}{p_1 + p_2 + (p_1 h_2 + p_2 h_1)t} - 1 \right] \frac{1}{t}.$$

Since  $p_1 + p_2 = 1$  we obtain:

$$\begin{aligned} h(t) &= \left[ \frac{1 + (h_1 + h_2)t + h_1 h_2 t^2}{1 + (p_1 h_2 + p_2 h_1)t} - 1 \right] \frac{1}{t} = \frac{(h_1 + h_2 - p_1 h_2 - p_2 h_1)t + h_1 h_2 t^2}{1 + (p_1 h_2 + p_2 h_1)t} \cdot \frac{1}{t} \\ &= \frac{h_1 + h_2 - p_1 h_2 - p_2 h_1 + h_1 h_2 t}{1 + (p_1 h_2 + p_2 h_1)t} = \frac{p_1 h_1 + p_2 h_2 + h_1 h_2 t}{1 + (p_1 h_2 + p_2 h_1)t}. \end{aligned}$$

By differentiating  $h(t)$ :

$$h'(t) = \frac{h_1 h_2 (1 + (p_1 h_2 + p_2 h_1)t) - (p_1 h_1 + p_2 h_2 + h_1 h_2 t)(p_1 h_2 + p_2 h_1)}{[1 + (p_1 h_2 + p_2 h_1)t]^2} \quad (13)$$



We need to show that  $h'(t) < 0$ . Since the denominator of (13) is positive, the sign of  $h'(t)$  depends on the sign of the numerator. Therefore, we denote the numerator of (13) by  $Q$  and analyse it separately:

$$\begin{aligned} Q(t) &= h_1 h_2 [1 + (p_1 h_2 + p_2 h_1) t] - (p_1 h_1 + p_2 h_2 + h_1 h_2 t) (p_1 h_2 + p_2 h_1) \\ &= h_1 h_2 + h_1 h_2 (p_1 h_2 + p_2 h_1) t - (p_1 h_1 + p_2 h_2) (p_1 h_2 + p_2 h_1) - h_1 h_2 (p_1 h_2 + p_2 h_1) t \\ &= h_1 h_2 - (p_1 h_1 + p_2 h_2) (p_1 h_2 + p_2 h_1). \end{aligned}$$

By expanding the brackets and using the fact that  $p_1 + p_2 = 1$  implies  $1 - p_1^2 - p_2^2 = 2p_1 p_2$  expression  $Q$  can be simplified further:

$$\begin{aligned} Q(t) &= h_1 h_2 - p_1^2 h_1 h_2 - p_1 p_2 h_1^2 - p_1 p_2 h_2^2 - p_2^2 h_1 h_2 \\ &= h_1 h_2 (1 - p_1^2 - p_2^2) - p_1 p_2 (h_1^2 + h_2^2) \\ &= 2p_1 p_2 h_1 h_2 - p_1 p_2 (h_1^2 + h_2^2) \\ &= -p_1 p_2 (h_1 - h_2)^2. \end{aligned}$$

Therefore, since  $h_1 \neq h_2$  we have  $Q < 0$ . Hence it follows that  $h'(t) < 0$  and  $h(t)$  is strictly decreasing.

Suppose that the proposition holds for  $n = k$ . We need to show that it also holds for  $n = k + 1$ . When  $n = k + 1$  the certainty equivalent hyperbolic discount rate is:

$$h_{k+1}(t) = \left[ \frac{1}{D^{(k+1)}} - 1 \right] \frac{1}{t},$$

where

$$D^{(k+1)} = \sum_{i=1}^{k+1} p_i D_i = (1 - p_{k+1}) \left( \sum_{i=1}^k \frac{p_i}{1 - p_{k+1}} D_i \right) + p_{k+1} D_{k+1}.$$

Since

$$\sum_{i=1}^k \frac{p_i}{1 - p_{k+1}} = 1,$$

we have

$$D^{(k+1)} = (1 - p_{k+1}) D^{(k)} + p_{k+1} D_{k+1}.$$

where

$$D^{(k)} = \sum_{i=1}^k \frac{p_i}{1 - p_{k+1}} D_i.$$

By the induction hypothesis it follows that

$$D^{(k)} = \frac{1}{1 + h_k(t)t},$$

where  $h_k$  is strictly decreasing. Therefore,

$$\begin{aligned} h^{(k+1)}(t) &= \left[ \frac{1}{(1 - p_{k+1}) D^{(k)} + p_{k+1} D_{k+1}} - 1 \right] \frac{1}{t} \\ &= \left[ \frac{1}{(1 - p_{k+1}) (1 + h_k(t)t)^{-1} + p_{k+1} (1 + h_{k+1}t)^{-1}} - 1 \right] \frac{1}{t}. \end{aligned}$$

Let  $\hat{p}_1 = 1 - p_{k+1}$ ,  $\hat{p}_2 = p_{k+1}$ ,  $\hat{h}_1(t) = h_k(t)$  and  $\hat{h}_2 = h_{k+1} = \text{const.}$  Then we have

$$h^{(k+1)}(t) = \left[ \frac{1}{\hat{p}_1(1 + \hat{h}_1(t)t)^{-1} + \hat{p}_2(1 + \hat{h}_2t)^{-1}} - 1 \right] \frac{1}{t}.$$

Analogously to the case  $n = 2$ , this expression can be rearranged to give:

$$h^{(k+1)}(t) = \frac{\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t}{1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t}.$$

However, by contrast to the case  $n = 2$ ,  $\hat{h}_1$  is now a function of  $t$ . The derivative of  $h^{(k+1)}$  is:

$$\frac{dh^{(k+1)}(t)}{dt} = \frac{\left(\hat{p}_1\hat{h}_1' + \hat{h}_1\hat{h}_2 + \hat{h}_1'\hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right)\left(\hat{p}_1\hat{h}_2 + \hat{p}_2\hat{h}_1 + \hat{p}_2\hat{h}_1't\right)}{\left[1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right]^2}.$$

The denominator of this fraction is strictly positive, so the sign of the derivative depends on the numerator only. Denote the numerator by  $N$ :

$$N = \left(\hat{p}_1\hat{h}_1' + \hat{h}_1\hat{h}_2 + \hat{h}_1'\hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right)\left(\hat{p}_1\hat{h}_2 + \hat{p}_2\hat{h}_1 + \hat{p}_2\hat{h}_1't\right).$$

Note that

$$N = \hat{Q}(t) + \hat{h}_1' \left[ \left(\hat{p}_1 + \hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \hat{p}_2t\left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right) \right],$$

where  $\hat{Q}(t)$  is defined as in the proof of Proposition 1, but with  $h_1 = \hat{h}_1(t)$  and  $h_2 = \hat{h}_2$ . Since Proposition 1 establishes that  $\hat{Q}(t) \leq 0$  (with equality if and only if  $\hat{h}(t) = h_2$ ) and  $\hat{h}_1' < 0$ , it suffices to show that

$$\left(\hat{p}_1 + \hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \hat{p}_2t\left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right) > 0 \quad (14)$$

Cancelling terms on the left-hand side of (14) leaves us with:

$$\hat{p}_1\left(1 + \hat{p}_1\hat{h}_2t\right) + \hat{h}_2t\left(1 + \hat{p}_1\hat{h}_2t\right) - (\hat{p}_2)^2\hat{h}_2t.$$

We now use the fact that  $(\hat{p}_2)^2 = (1 - \hat{p}_1)^2 = 1 - 2\hat{p}_1 + (\hat{p}_1)^2$  to get

$$\begin{aligned} & \hat{p}_1\left(1 + \hat{p}_1\hat{h}_2t\right) + \hat{h}_2t\left(1 + \hat{p}_1\hat{h}_2t\right) - [1 - 2\hat{p}_1 + (\hat{p}_1)^2]\hat{h}_2t \\ &= \hat{p}_1 + \left(\hat{h}_2t\right)^2\hat{p}_1 + 2\hat{p}_1\hat{h}_2t, \end{aligned}$$

which is strictly positive as required. Therefore,  $h^{(k+1)}(t)$  is strictly decreasing.  $\square$

## 5 Discussion

We generalized Jackson and Yariv’s result [11] by proving that whenever we aggregate different discount functions from comparable DI classes, the weighted average function is always *strictly more DI* than the least DI of its constituents. This also strengthens the conclusion of the theorem of Prelec [19] who demonstrates that the mixture of two different discount functions from the same DI class represents more DI preferences.

When a decision maker is uncertain about her hyperbolic discount rate, we showed that long-term costs and benefits must be discounted at the probability-weighted harmonic mean of the hyperbolic discount rates that might eventuate. This complements the well-known result of Weitzman [27].

One natural question that arises is whether it is possible to prove a result analogous to Proposition 13 when all preference orders exhibit increasing impatience (II). Will the mixture of II discount functions be (strictly) II? Perhaps surprisingly, the answer to this question is negative in general.

This follows from results in the literature on survival analysis and reliability theory. The similarity between reliability theory and temporal discounting is discussed in [24]. Takeuchi [26] also notes that a discount function is analogous to a survival function,  $S(t)$ . The failure rate associated with  $S(t)$  is

$$g(t) = -\frac{S'(t)}{S(t)},$$

which behaves as a time preference rate. For twice continuously differentiable survival functions, a decreasing failure rate (DFR) corresponds to a decreasing time preference rate, and hence to DI, whereas an increasing failure rate (IFR) corresponds to II. Mixtures of probability distributions are a common topic in survival and reliability analysis. Proschan [21] established that mixtures of distributions with DFR always exhibit DFR.<sup>6</sup> However, Gurland and Sethuraman [8, 9] provide striking examples of mixtures of very quickly increasing failure rates that are eventually decreasing.

## 6 Appendix

### 6.1 Fishburn and Rubinstein’s axioms for a discounted utility representation

After Fishburn and Rubinstein [7], we assume that:

**Axiom 1. (Weak Order)** The preference order  $\succsim$  is a weak order, i.e., it is complete and transitive.

**Axiom 2. (Monotonicity)** For every  $x, y \in X$ , if  $x < y$ , then  $(x, t) \prec (y, t)$  for every  $t \in T$ .

---

<sup>6</sup>This result is comparable to the “non-strict” part of our Proposition 13.

**Axiom 3. (Continuity)** For every  $(y, s) \in X \times T$  the sets  $\{(x, t) \in X \times T: (x, t) \succ (y, s)\}$  and  $\{(x, t) \in X \times T: (x, t) \preccurlyeq (y, s)\}$  are closed.

**Axiom 4. (Impatience)** For all  $t, s \in T$  and every  $x > 0$ , if  $t < s$ , then  $(x, t) \succ (x, s)$ .  
If  $t < s$  and  $x = 0$ , then  $(x, t) \sim (x, s)$  for every  $t, s \in T$ , that is, 0 is a time-neutral outcome.

**Axiom 5. (Separability)** For every  $x, y, z \in X$  and every  $r, s, t \in T$  if  $(x, t) \sim (y, s)$  and  $(y, r) \sim (z, t)$  then  $(x, r) \sim (z, s)$ .

Fishburn and Rubinstein [7] proved the following result:

**Theorem 4** ([7]). *The preferences  $\succcurlyeq$  on  $X \times T$  satisfy Axioms 1-5 if and only if there exists a discounted utility representation for  $\succcurlyeq$  on  $X \times T$ . If  $(u, D)$  and  $(u_0, D_0)$  both provide discounted utility representations for  $\succcurlyeq$  on  $X \times T$ , then  $u = \alpha u_0$  for some  $\alpha > 0$ , and  $D = \beta D_0$  for some  $\beta > 0$ .*

## 6.2 Proof of Proposition 7

We need to prove the following lemma first:

**Lemma 18.** *Suppose that  $h_1$  and  $h_2$  are strictly decreasing functions. Then  $h_1$  is a (strictly) convex transformation of  $h_2$  if and only if  $h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma)$  implies that  $h_1(s) - h_1(t) \leq [\prec] h_1(s + \sigma + \rho) - h_1(t + \sigma)$  for every  $s, t, \sigma$  and  $\rho$  satisfying  $0 < t < s \leq t + \sigma < s + \sigma + \rho$ .*

*Proof.* We prove necessity first. Suppose that  $h_1$  is a (strictly) convex transformation of  $h_2$ ; that is, there exists a (strictly) convex function  $f$  such that  $h_1 = f(h_2)$ . Assume also that  $0 < t < s \leq t + \sigma < s + \sigma + \rho$  and

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma). \quad (15)$$

We need to show that

$$h_1(s) - h_1(t) \leq [\prec] h_1(s + \sigma + \rho) - h_1(t + \sigma)$$

whenever  $0 < t < s \leq t + \sigma < s + \sigma + \rho$ . Since  $h_2$  is strictly decreasing, it follows that

$$h_2(s + \sigma + \rho) < h_2(t + \sigma) \leq h_2(s) < h_2(t).$$

Recall that  $f$  is a (strictly) convex function. Therefore, as equality (15) holds, it implies that

$$f(h_2(t + \sigma)) - f(h_2(s + \sigma + \rho)) \leq [\prec] f(h_2(t)) - f(h_2(s)).$$

Since  $h_1 = f(h_2)$ , this inequality is equivalent to

$$h_1(t + \sigma) - h_1(s + \sigma + \rho) \leq [\prec] h_1(t) - h_1(s).$$

Rewriting:

$$h_1(s) - h_1(t) \leq [ < ] h_1(s + \sigma + \rho) - h_1(t + \sigma), \quad (16)$$

whenever  $0 < t < s \leq t + \sigma < s + \sigma + \rho$ .

To show the sufficiency, suppose that (15) implies (16) for every  $s, t, \sigma$  and  $\rho$  satisfying  $0 < t < s \leq t + \sigma < s + \sigma + \rho$ . Define  $f$  such that  $f = h_1 \circ h_2^{-1}$ . Note that we can do so because  $h_2^{-1}$  exists (since  $h_2$  is a strictly decreasing function). Then if

$$h_2(s + \sigma + \rho) < h_2(t + \sigma) \leq h_2(s) < h_2(t)$$

and equation (15) holds, we have

$$f(h_2(t + \sigma)) - f(h_2(s + \sigma + \rho)) \leq [ < ] f(h_2(t)) - f(h_2(s)).$$

Therefore,  $f$  is a (strictly) convex function, which means that  $h_1$  is a (strictly) convex transformation of  $h_2$ .  $\square$

We can now prove Proposition 7.

*Proof.* Observe that  $D_i: [0, \infty) \rightarrow (0, 1]$  is one-to-one and onto, so  $D_i^{-1}: (0, 1] \rightarrow [0, \infty)$ .

Let us first prove that condition (i) follows from condition (ii). The proof is by contraposition. We show that not (i) implies not (ii). Assume that (i) fails; that is, there exist  $s$  and  $t$  with  $0 < t < s$ ,  $\rho > 0$ ,  $\sigma > 0$  and  $x, y, x', y' \in X$  with  $0 < x < y$  and  $0 < x' < y'$  such that  $(x', t) \sim_2 (y', s)$ ,  $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$ ,  $(x, t) \sim_1 (y, s)$  and

$$(x, t + \sigma) \succ_1 [\succ_1] (y, s + \sigma + \rho).$$

Since  $u_1(y) > 0$  and  $u_2(y') > 0$  by assumption, this implies

$$\frac{u_2(x')}{u_2(y')} = \frac{D_2(s)}{D_2(t)} = \frac{D_2(s + \sigma + \rho)}{D_2(t + \sigma)}$$

and

$$\frac{u_1(x)}{u_1(y)} = \frac{D_1(s)}{D_1(t)} > [\geq] \frac{D_1(s + \sigma + \rho)}{D_1(t + \sigma)}.$$

Let  $h_1 = \ln D_1$  and  $h_2 = \ln D_2$ . Note that  $h_1$  and  $h_2$  are both strictly decreasing functions. Observe also that  $h_i: [0, \infty) \rightarrow (-\infty, 0]$  is one-to-one and onto. Thus  $h_i^{-1}: (-\infty, 0] \rightarrow [0, \infty)$ , where  $h_i^{-1}(z) = D_i^{-1}(e^z)$ . Rewriting these expressions we get  $D_i(t) = e^{h_i(t)}$  for each  $i \in \{1, 2\}$ . Thus:

$$\frac{e^{h_2(s)}}{e^{h_2(t)}} = \frac{e^{h_2(s + \sigma + \rho)}}{e^{h_2(t + \sigma)}}$$

and

$$\frac{e^{h_1(s)}}{e^{h_1(t)}} > [\geq] \frac{e^{h_1(s + \sigma + \rho)}}{e^{h_1(t + \sigma)}}.$$

Equivalently,

$$h_2(s) - h_2(t) = h_2(s + \rho + \sigma) - h_2(t + \sigma) \quad (17)$$

and

$$h_1(s) - h_1(t) > [\geq] h_1(s + \rho + \sigma) - h_1(t + \sigma). \quad (18)$$

Note that  $\ln D_1(D_2^{-1}(e^z))$  (strictly) convex in  $z$  on  $(-\infty, 0]$  is equivalent to  $h_1 \circ h_2^{-1}$  (strictly) convex in  $z$  on  $(-\infty, 0]$ . In other words,  $h_1$  is a (strictly) convex transformation of  $h_2$ . By Lemma 18 this conclusion contradicts equation (17) and inequality (18). Therefore, not (i) implies not (ii).

Secondly, we need to demonstrate that (i) implies (ii). Using the previously introduced notation, we show that for every  $s, t, \sigma$  and  $\rho$  satisfying

$$0 < t < s \leq t + \sigma < s + \sigma + \rho$$

the equation

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma)$$

implies

$$h_1(s) - h_1(t) \leq [<] h_1(s + \sigma + \rho) - h_1(t + \sigma).$$

As  $h_1$  and  $h_2$  are decreasing functions, this proves that  $h_1$  is a (strictly) convex transformation of  $h_2$ . Assume that  $0 \leq t < s \leq t + \sigma < s + \sigma + \rho$  such that

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma).$$

By definition of  $h_i = \ln D_i$  this expression is equivalent to

$$\frac{D_2(s)}{D_2(t)} = \frac{D_2(s + \sigma + \rho)}{D_2(t + \sigma)} \in (0, 1).$$

As  $u_2$  is continuous, we can choose  $0 < x' < y'$  such that:

$$\frac{D_2(s)}{D_2(t)} = \frac{D_2(s + \sigma + \rho)}{D_2(t + \sigma)} = \frac{u_2(x')}{u_2(y')}.$$

Therefore,  $D_2(t)u_2(x') = D_2(s)u_2(y')$  and  $D_2(t + \sigma)u_2(x') = D_2(s + \sigma + \rho)u_2(y')$ . This means that  $(x', t) \sim_2 (y', s)$  and  $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$ .

Analogously, because  $u_1$  is continuous, we can choose  $x, y$  such that:

$$\frac{D_1(s)}{D_1(t)} = \frac{u_1(x)}{u_1(y)} \in (0, 1).$$

Hence,  $(x, t) \sim_1 (y, s)$ .

But according to (i), if  $(x', t) \sim_2 (y', s)$ ,  $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$  and  $(x, t) \sim_1 (y, s)$  then  $(x, t + \sigma) \preceq_1 [\prec_1] (y, s + \sigma + \rho)$ . The latter is equivalent to:

$$\frac{D_1(s + \sigma + \rho)}{D_1(t + \sigma)} \geq [>] \frac{u_1(x)}{u_1(y)}.$$

It follows that

$$\frac{D_1(s)}{D_1(t)} \leq [<] \frac{D_1(s + \sigma + \rho)}{D_1(t + \sigma)},$$

which is equivalent to

$$\ln D_1(s) - \ln D_1(t) \leq [ < ] \ln D_1(s + \sigma + \rho) - \ln D_1(t + \sigma)$$

or

$$h_1(s) - h_1(t) \leq [ < ] h_1(s + \sigma + \rho) - h_1(t + \sigma).$$

Therefore,

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma)$$

implies

$$h_1(s) - h_1(t) \leq [ < ] h_1(s + \sigma + \rho) - h_1(t + \sigma)$$

whenever  $0 \leq t < s \leq t + \sigma < s + \sigma + \rho$ . Hence, by Lemma 18,  $h_1$  is a (strictly) convex transformation of  $h_2$ .  $\square$

## References

- [1] G. Ainslie. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4):463–496, 1975.
- [2] A. Al-Nowaihi and S. Dhami. A note on the Loewenstein-Prelec theory of intertemporal choice. *Mathematical Social Sciences*, 52(1):99–108, 2006.
- [3] K. J. Arrow. *Aspects of the theory of risk-bearing*. Yrjö Jahnssonin Säätiö, Helsinki, 1965.
- [4] A. E. Attema, H. Bleichrodt, K. I. M. Rohde, and P. P. Wakker. Time-tradeoff sequences for analyzing discounting and time inconsistency. *Management Science*, 56(11):2015–2030, 2010.
- [5] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.
- [6] Z. Cvetkovski. *Inequalities: Theorems, Techniques and Selected Problems*. Springer-Verlag, Berlin Heidelberg, 2012.
- [7] P. C. Fishburn and A. Rubinstein. Time preference. *International Economic Review*, 23(3):677–694, 1982.
- [8] J. Gurland and J. Sethuraman. Shorter communication: reversal of increasing failure rates when pooling failure data. *Technometrics*, 36(4):416–418, 1994.
- [9] J. Gurland and J. Sethuraman. How pooling failure data may reverse increasing failure rates. *Journal of the American Statistical Association*, 90(432):1416–1423, 1995.
- [10] C. M. Harvey. Proportional discounting of future costs and benefits. *Mathematics of Operations Research*, 20(2):381–399, 1995.

- [11] M. O. Jackson and L. Yariv. Present bias and collective dynamic choice in the lab. *the American Economic Review*, 104(12):4184–4204, 2014.
- [12] M. O. Jackson and L. Yariv. Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–178, 2015.
- [13] D. T. Jamison and J. Jamison. Characterizing the amount and speed of discounting procedures. *Journal of Benefit-Cost Analysis*, 2(2):1–56, 2011.
- [14] D. Laibson. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–478, 1997.
- [15] G. Loewenstein and D. Prelec. Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597, 1992.
- [16] J. E. Mazur. Hyperbolic value addition and general models of animal choice. *Psychological Review*, 108(1):96–112, 2001.
- [17] E. S. Phelps and R. A. Pollak. On second-best national saving game-equilibrium growth. *Review of Economic Studies*, 35(2):185–199, 1968.
- [18] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964.
- [19] D. Prelec. Decreasing impatience: A criterion for non-stationary time preference and hyperbolic discounting. *Scandinavian Journal of Economics*, 106(3):511–532, 2004.
- [20] D. Prelec and G. Loewenstein. Decision making over time and under uncertainty: A common approach. *Management Science*, 37(7):770–786, 1991.
- [21] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3):373–383, 1963.
- [22] J. Quiggin and J. Horowitz. Time and risk. *Journal of Risk and Uncertainty*, 10(1):37–55, 1995.
- [23] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag, Berlin Heidelberg, 1998.
- [24] P. D. Sozou. On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1409):2015–2020, 1998.
- [25] O. Stein. Twice differentiable characterizations of convexity notions for functions on full dimensional convex sets. *Schedae Informaticae*, 21:55–63, 2012.
- [26] K. Takeuchi. Non-parametric test of time consistency: Present bias and future bias. *Games and Economic Behavior*, 71(2):456–478, 2011.



- [27] M. L. Weitzman. Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management*, 36(3):201–208, 1998.